

The ISO Standards are powerful tools for describing many aspects of environmental data. As is so often the case, powerful tools seem complex or difficult to use at the start. The goal of this wiki is to create a place where the community can help itself ask and answer questions clearly. This page offers might be a useful place to start. Please feel free to post questions here or to email them to ted.habermann@noaa.gov who will do his best to start the process of answering them!

Contents

- 1 General Questions
 - ◆ 1.1 What is covered by the ISO Metadata Standards?
 - ◆ 1.2 What are all of these numbers?
 - ◆ 1.3 What are the ISO metadata pictures in this wiki?
 - ◆ 1.4 What Costs Are Associated With ISO Documents?
 - ◆ 1.5 What Costs Are Associated With Translating from FGDC to ISO?
 - ◆ 1.6 What are the benefits to organizations and data/metadata creators of switching from CSDGM to ISO metadata?
 - ◆ 1.7 How often will the ISO standards change?
 - ◆ 1.8 What software tools are now available for working with ISO metadata?
- 2 Using the ISO Standard
 - ◆ 2.1 What is the difference between MD Metadata and MI Metadata?
 - ◆ 2.2 How do I decide whether to use the MD Package or MI Package?
 - ◆ 2.3 Where would you put a Digital Object Identifier (DOI) in the metadata?
 - ◆ 2.4 What is the Dataset URI?
 - ◆ 2.5 How do we describe datasets and services together?
- 3 ISO XML
 - ◆ 3.1 What is XML?
 - ◆ 3.2 How do I get started writing ISO XML?
 - ◆ 3.3 Can you show examples of validation errors and resolutions for ISO records?
 - ◆ 3.4 What are MD , MI , DS , ...?
 - ◆ 3.5 What are gmd:, gmi:, gml:, ...?
 - ◆ 3.6 What is an ISO code list and How can ISO code lists be extended?
 - ◆ 3.7 What if I want to constrain a free text field?

General Questions

What is covered by the ISO Metadata Standards?



ISO Objects.png

This Figure shows the principle objects included in the ISO Metadata Standard and whether they are contained by MI/MD_Metadata or by MD_DataIdentification. Pages in the [ISO 19115 Category](#) give more information for many of these objects.

What are all of these numbers?

ISO Standards are identified using numbers like 19115. ISO Technical Committee 211 Geographic information/Geomatics ([TC211](#)) is responsible for the ISO geographic information series of standards. The most commonly used numbers in our discussion are:

Standard	Topic	Wiki Category
ISO 19115	Geographic information ? Metadata	Category
ISO 19115-2	Geographic information ? Metadata - Part 2: Extensions for imagery and gridded data	
ISO 19119	Metadata for services	Category
ISO 19139	Metadata - XML schema implementation	Category

More numbers and fact sheets are available in the TC211 [Standards Guide](#) or at the [TC211 Home Page](#).

Note: Unfortunately, the name of 19115-2 is a bit misleading. It is essentially an extension of 19115 that adds information about missions, platforms and sensors, along with a few other things. Most environmental observations involve sensors, so 19115-2 is probably the right standard to use for most environmental datasets, not just imagery or gridded data.

What are the ISO metadata pictures in this wiki?

The ISO Standards, and many others, are being developed and described using the [Unified Modeling Language](#) (UML), a standardized general-purpose modeling language from the field of software engineering. Most of the pictures in the wiki are UML [Class Diagrams](#) that describe the structure of a metadata by showing the classes, their attributes, and the relationships between the classes.

All of the official ISO 19115 UML diagrams are available from [ISO](#) at no cost as part of the correction to ISO 19115. That document includes the pictures, but not the data dictionary which defines the elements. In many cases the pictures are enough.

What Costs Are Associated With ISO Documents?

ISO supports the development of standards through sales of the finished documents that are available at the [ANSI Store](#). These are technical documents that contain pictures of the UML Models and definitions of the elements along with some, but generally not much, explanatory material. The [XML Schemas](#) are available at no cost.

This wiki includes over 250 pages that discuss many aspects of the ISO standards along with examples and experiences. It is developed by a community of practitioners from the real-world. Others in the community can contribute their examples and experiences as we move forward. This is a free and dynamic resource that is likely to be more useful in many ways than the actual standards documents. In addition, you can contribute and make it even better!

What Costs Are Associated With Translating from FGDC to ISO?

Significant costs associated with development and testing of the FGDC to ISO [transforms](#) have already been taken care of by NCDDC and NGDC. This work has been going on for several years and many, if not most, of the rough spots in the road have been grappled with and smoothed over. The transforms can be implemented on the desktop using standard XML editors, using NCDDC's MERMAid tool, and NGDC has

tools for automatically transforming entire collections. In addition, over 250 pages have been created in this wiki that discuss many aspects of the ISO standards along with examples and experiences. This wiki will continue to grow as more contributors gain experience and create helpful examples. This open knowledge sharing environment makes it possible for Federal Agencies and groups around the world to join the ISO community with very low start-up costs. Once the translation is done, the maintenance costs should be roughly equivalent to the cost of maintaining the original FGDC metadata.

What are the benefits to organizations and data/metadata creators of switching from CSDGM to ISO metadata?

There are at least two groups of benefits to switching from CSDGM to ISO. The first group is related to the worldwide community of ISO users. It is becoming more and more important to share many datasets across international borders. Use and understanding of these data is clearly facilitated by using international documentation standards. The second group of benefits are related to the much improved capabilities of the ISO standards both because of the sophisticated use of XML and because of the breadth of the standard. See [Use Cases to CRUD](#) and [new capabilities](#) for a discussion of some of these benefits.

How often will the ISO standards change?

ISO Technical Committee 211 (TC211) produces two types of documents: International Standards and Technical Specifications. International Standards are eligible for revision every five years. When a standard becomes eligible for revision, members of TC211 (including the U.S.) vote and a revision happens only if they vote for one. Technical Specifications are easier to create and last for three years at which time they become a standard or expire.

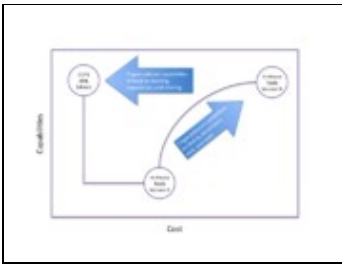
ISO 19115 was created in 2003 and became eligible for revision during 2008. Members supported a revision and an editing committee was formed. The task of the editing committee was finished during fall of 2011 with the acceptance of ISO 19115-1 as a Draft International Standard (DIS). That draft is now out for review and comment and will likely have some editorial and minor revisions during the summer of 2012. The current plan is for it to become a Final Draft International Standard (FDIS) during November, 2012 and an International Standard (IS) during May, 2013. The periods between these phases are determined by the ISO rules and they exist so that members can fully evaluate the versions before voting.

In this real-world case, the original ISO 19115 was stable for ten years following its adoption as an International Standard during 2003. The revision (19115-1) is now a Draft International Standard and many groups are comfortable starting to experiment with the implementation as only minor technical changes are expected. If the standard also lasts ten years prior to revision, it will be stable until 2023.

It is important to understand that the revisions to 19115 are driven by real requirements of the members that participate in the ISO process. In this case they significantly improve the [capabilities](#) of the standard. The same is true for recent revisions of the data quality standards (see [ISO 19157](#)). In these cases, change is definitely good!

What software tools are now available for working with ISO metadata?

The [Metadata Tools](#) Category includes information about a variety of tools for working with ISO metadata and other XML dialects.



Capabilities and Costs

The ISO tool landscape is very different than that which existed for CSDGM because of the incredible development that has occurred in the world of XML over the last decade. The ISO standards build on these developments to bring a rich set of capabilities to metadata creators and users. The most capable tools for working with ISO metadata are full featured XML editors like Oxygen and XMLSpy. These tools include all of the necessary capabilities including schema-aware editing, validation using schema or schematron, transformations using XSLT, remote access to files and transforms, tagless editing, and batch processing.

As illustrated in the Figure to the right, these very capable tools are available today at very low costs. Despite that, it is common to hear people say that they are still waiting for ISO tools. The implication of that statement is that they are waiting for someone to develop these tools. This is the classic "build vs. buy" question. In this case, experience suggests that the build option is immediately more expensive and, because very capable tools already exist, results in an immediate decrease in capabilities. Overcoming that capability decrease is time consuming and expensive and comes with a long-term maintenance burden which adds significant costs.

The cost question is important, but the more interesting question involves how organizational capabilities are increased in these two scenarios. In the build scenarios, capability increases are limited by funding, developer skills, and the development/testing/deployment cycle. Users communicate needs to developers and work with them to make sure they get it right. This takes time and inevitably delays the introduction of the new capability.

In the buy scenario, all of the necessary capabilities already exist in the tools and their adoption across an organization depends on training, experience, and sharing of skills and knowledge. This wiki is an example of how a community can increase capabilities using this approach. For example, there are many pages with information on using oxygen to perform many common tasks. Members of the community that are familiar with other editors and tools are encouraged to provide similar information for those tools and all are welcome to contribute experiences and examples as we move forward.

Using the ISO Standard

What is the difference between MD Metadata and MI Metadata?



MD Metadata.png

ISO_FAQ

The ISO MD Metadata Object is the container for all other 19115 metadata elements. The elements included in the MD_Metadata container generally describe the metadata rather than the dataset.

The MD_Metadata object was extended in Part 2 of the ISO Metadata Standard (19115-2) to include the MI AcquisitionInformation object (shaded in this Figure) for describing platforms, instruments, and other aspects of data acquisition. This extension requires changing the name and the namespace of the object. The new name is MI_Metadata and the namespace is gmi (root element is gmi:MI_Metadata instead of gmd:MD_Metadata).

Most environmental datasets start with observations made using instruments, and knowledge of the instrument characteristics can be critical to understanding the data. Therefore, it makes sense to use the full ISO Standard (19115 + 19115-2) for documenting most environmental data.

How do I decide whether to use the MD Package or MI Package?

The MI Package is a super-set of the MD Package that can provide important additional metadata, like descriptions of the instrumentation used to make the measurements and the processing of those measurements. The recommended practice is to use MI even if you do not currently have any of this additional information.

The MI Package is now supported by geo.data.gov and Geoportal.

Where would you put a Digital Object Identifier (DOI) in the metadata?

Digital Object Identifiers (DOI) are most commonly used to identify and cite published datasets. In the ISO standard these identifiers would be included as an MD_Identifier in the CI Citation for the dataset. Like in the FGDC, this citation describes how the dataset that the metadata describes is to be cited. That citation would include the DOI for the dataset. If the metadata record itself also had a DOI, that would be in the fileIdentifier (more).

What is the Dataset URI?

The datasetURI is a Universal Resource Identifier (URI) for the dataset. URIs have been around for a long time but the problem of how they get resolved into the resource they identify has never been solved. In other words, once you have a URI - what do you do with it to find the resource? The dataSetURI has been dropped in the revision of 19115 because of this confusion. If you have a unique identifier for the dataset, it goes in the MD_Identifier in the CI_Citation for the dataset (see DOI discussion above). If you have a Universal Resource Locator (URL) for the dataset, it goes in the new CI_OnlineResource that is included in the revised CI_Citation.

How do we describe datasets and services together?

The ISO standard allows data and the services that serve them to be described in the same record. This approach is used extensively in the ncISO capability for the THREDDS Data Server which routinely makes datasets available through multiple services. See the IOOS records for examples. Those records use the distribution information to describe access methods designed for humans and service identification for services designed for machine or application access.

An alternative approach involves linking dataset records to service records. This is not encouraged as there is no clearly defined mechanism for implementing or using such links.

What is the difference between MD_Metadata and MI_Metadata?

ISO XML

What is XML?

XML (Extensible Markup Language) is a standard way of formatting information so that machines can read and understand it. There are many books and tutorials available. This [tutorial](#) has been recommended. The XML is basically made up of tags, elements (with or without content) and attributes.

A **Tag** is a markup construct that begins with "<" and ends with ">". Closing (ending) tags will always contain "/".

Ex:

start tags <tagName>

end tags </tagName>

Elements are a logical component of a document which either begins with a start-tag and ends with a matching end-tag, or consists only of an empty-element tag. Any characters between the start- and end-tags are the element's content, and may contain markup, including other elements, which are called child elements. In the example below, the element is named "Greeting" and the content of the element is "Hello, world."

Ex:

<Greeting>Hello, world.</Greeting>

Attributes are a markup construct consisting of a name/value pair that exists within a start-tag or empty-element tag. More information on the attributes used in ISO are found [here](#).

How do I get started writing ISO XML?

ISO XML is robust and complex. Some recommendations:

- *Do not start writing ISO from scratch.* Start with an example that is similar to what you need. Many examples are available from [NGDC](#). The authors of those examples have made many mistakes that you do not want to repeat!
- Use a complete XML editor with schema, validation, and transform support. [Altova XMLSpy](#) and [Oxygen](#) both have free 30-day trials. Some [Oxygen Tips](#) are available.
- Use a schema to validate the ISO early and often.
- If you have existing FGDC or DIF metadata, use a [translator](#) to start your ISO record.
- [ISO Boilerplate](#) provides descriptions of content used in all ISO records.
- [Documentation Spirals](#) break the ISO Standard into smaller pieces designed to be easy to understand and implement. They can be used as a roadmap for building complete documentation. The [Rubric Transform](#) can be used with an XML editor to facilitate that process.
- Explore this wiki for information and examples.
- Contribute questions and answers to this page.

Can you show examples of validation errors and resolutions for ISO records?

Information about validation procedures and common errors is included in the [validation](#) category.

What are MD_, MI_, DS_, ...?

ISO elements are arranged into a number of groups or packages. Two letter abbreviations are used to denote the package that contains a class. Those abbreviations precede class names, connected by a `?`. A list of those abbreviations used in ISO follows.

- CI Citation (ISO 19115)
- DQ Data quality (ISO 19115)
- DS Dataset (ISO 19115)
- EX Extent (ISO 19115)
- FC Feature Catalogue (ISO 19110)
- GM Geometry (ISO 19115)
- LE Lineage Extended (ISO 19115-2)
- LI Lineage (ISO 19115)
- MD Metadata (ISO 19115)
- MI Imagery (ISO 19115-2)
- PT Polylinguistic Text (ISO 19115)
- QE Data Quality Extended (ISO 19115-2)
- RS Reference System (ISO 19115)
- SV Services (ISO 19119)

What are gmd:, gmi:, gml:, ...?

The ISO Metadata Standards are content standards; they define the content that can exist in metadata without information about how that content should be represented. They define semantics (meaning) of the metadata, not syntax (format). The most commonly used format, or representation, for ISO content is XML, a textual format which involves elements identified using angle brackets (tags) and content within those elements (the `/` in the second tag indicates an end tag): `<elementName>content</elementName>`. It is possible that the same `elementName` is used in multiple contexts or by multiple communities. XML includes the concept of namespaces in order to clarify the meaning in those situations. XML with namespaces looks like: `<namespace:elementName>content</namespace:elementName>`. ISO metadata includes several namespaces. They are generally specified at the top of the files and included in all element names. For example this xml:

```
<gmd:individualName>
<gco:CharacterString>Ted Habermann</gco:CharacterString>
</gmd:individualName>
```

includes an `individualName` element from the `gmd` namespace and a `CharacterString` element from the `gco` namespace.

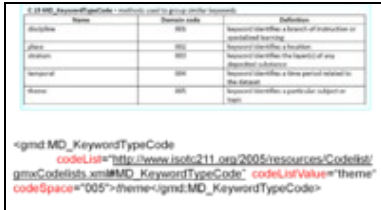
- See [ISO 19139 Identifiers](#) for more information about the general structure of ISO XML.
- See [ISO Namespaces](#) for more information about using namespaces in the ISO XML.

What is an ISO code list and How can ISO code lists be extended?

ISO_FAQ

In order to standardize allowable textual values for metadata elements, ISO Metadata makes use of code lists. A code list is an open enumeration of values. It is a flexible mechanism permitting to extend code lists as needed.

CodeLists are used in ISO metadata to provide a shared set of values for particular roles in the metadata. The CodeLists are used to refer to a specific codelist value in a register. Codelists contain the attributes *codeList*?, *codeListValue*?, and *codeSpace*?



The screenshot shows a table with columns: Name, Role, and Definition. It lists various ISO metadata elements and their roles. Below the table is an XML snippet for `<gmd:MD_KeywordTypeCode>` with attributes `codeList`, `codeListValue`, and `codeSpace`.

Name	Role	Definition
CharacterSet	MD	CharacterSet identifies a branch of instruction or associated learning.
Place	MD	Place identifies a location.
Medium	MD	Medium identifies the support of any associated information.
Temporal	MD	Temporal identifies a time period related to the content.
Theme	MD	Theme identifies a particular subject or topic.

```
<gmd:MD_KeywordTypeCode  
  codeList="http://www.isotc211.org/2005/resources/Codelist/  
gmx:Codelists.xml#MD_KeywordTypeCode" codeListValue="theme"  
  codeSpace="005">theme</gmd:MD_KeywordTypeCode>
```



CodeList.png

- The **codeList** attribute is *mandatory* and contains a URL that references a codeList definition within a registry or a codeList catalogue.
- The **codeListValue** attribute is also *mandatory* and contains the name of the selected value.
- The **codeSpace** attribute is *optional* and refers to the alternative expression of the codeListValue.

It is a best practice to also put the codeList value in the codeListValue attribute and in the element content:

```
<gmd:MD_CharacterSetCode codeList="http://www.isotc211.org/2005/resources/codeList.xml#MD_CharacterSetCode" codeListValue="Latin" codeSpace="005">Latin</gmd:MD_CharacterSetCode>
```

The codeLists were constructed this way in order to separate management of the codeList content from management of the standard. This means that metadata creators can supply your own codeList anywhere in the Standard by pointing to a custom location instead of the standard location. Of course, this could cause interoperability problems, so it should not be done without careful consideration of the implications.

NGDC codeList dictionaries can be found here: [ISO 19115 and 19115-2 CodeList Dictionaries](http://www.ngdc.noaa.gov/metadata/published/xsd/schema/resources/Codelist/gmx/Codelists.xml)

An example of the codeList dictionary is at

<http://www.ngdc.noaa.gov/metadata/published/xsd/schema/resources/Codelist/gmx/Codelists.xml>. BTW: the codeList elements are schema valid regardless of the actual content in either attribute. You need schematron validation in order to check the contents of the codeList elements.

What if I want to constrain a free text field?

In ISO metadata, you usually have either a `gco:CharacterString` allowing free text or a CodeList constraining terms to a registry. An anchor (**gmx:Anchor**) can be used as a substitute for a `CharacterString` when you want to restrict the free text to a controlled vocabulary.

Ex :

```
<gco:CharacterString>Pearl Harbor, HI</gco:CharacterString>
```


ISO_FAQ

can be replaced by

```
<gmx:Anchor xlink:href="http://www.rvdata.us/voc/port#101065">Pearl Harbor, HI</gmx:Anchor>
```

This restricts the field to the controlled vocabulary located at <http://www.rvdata.us/voc/port> and